



Entrez PubMed Nucleotide Protein Genome Structure PMC Taxonomy Boo

Search Protein for Go Clear

Limits Preview/Index History Clipboard Details  
Display default Show: 20 Send to File Get Subsequence Featur

1: Q16853. Reports Membrane copper a...

BLink, Domains,

[gi:2501336]

Links

LOCUS Q16853 763 aa linear PRI 15-JUN-2004

DEFINITION Membrane copper amine oxidase (Vascular adhesion protein-1) (VAP-1) (HPAO).

ACCESSION Q16853

VERSION Q16853 GI:2501336

DBSOURCE swissprot: locus AOC3\_HUMAN, accession Q16853;  
class: standard.  
created: Nov 1, 1997.  
sequence updated: Nov 1, 1997.  
annotation updated: Jun 15, 2004.  
xrefs: gi: [1399031](#), gi: [1399032](#), gi: [3283369](#), gi: [3283370](#), gi: [11602472](#), gi: [11602473](#), gi: [11602470](#), gi: [11602471](#), gi: [29792246](#), gi: [29792247](#), gi: [2134747](#)  
xrefs (non-sequence databases): GenewHGNC:550, MIM 603735, GO0016021, GO0005886, GO0008122, GO0005489, GO0009308, InterProIPR000269, PfamPF01179, PfamPF02727, PfamPF02728, PRINTSPR00766, PROSITEPS01164, PROSITEPS01165

KEYWORDS Oxidoreductase; Copper; TPQ; Glycoprotein; Transmembrane; Signal-anchor; Cell adhesion; Polymorphism; Metal-binding; Direct protein sequencing.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (residues 1 to 763)

AUTHORS Zhang,X. and McIntire,W.S.

TITLE Cloning and sequencing of a copper-containing, topa quinone-containing monoamine oxidase from human placenta

JOURNAL Gene 179 (2), 279-286 (1996)

MEDLINE 97128319

PUBMED 8972912

REMARK SEQUENCE FROM N.A.  
TISSUE=Placenta

REFERENCE 2 (residues 1 to 763)

AUTHORS Smith,D.J., Salmi,M., Bono,P., Hellman,J., Leu,T. and Jalkanen,S.

TITLE Cloning of vascular adhesion protein 1 reveals a novel multifunctional adhesion molecule

JOURNAL J. Exp. Med. 188 (1), 17-27 (1998)

MEDLINE 98317014

PUBMED 9653080

REMARK SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
TISSUE=Lung

REFERENCE 3 (residues 1 to 763)

AUTHORS Zhang,Q., Mashima,Y., Noda,S., Imamura,Y., Kudoh,J., Shimizu,N., Nishiyama,T., Umeda,S., Oguchi,Y., Tanaka,Y. and Iwata,T.

TITLE Characterization of AOC2 gene encoding a copper-binding amine oxidase expressed specifically in retina

JOURNAL Gene 318, 45-53 (2003)

MEDLINE 22948120  
 PUBMED 14585497  
 REMARK SEQUENCE FROM N.A.  
 REFERENCE 4 (residues 1 to 763)  
 AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,  
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,  
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,  
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,  
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,  
 Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
 Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,  
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,  
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
 TITLE Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 MEDLINE 22388257  
 PUBMED 12477932  
 REMARK SEQUENCE FROM N.A.  
 TISSUE=PNS  
 COMMENT -----  
 This SWISS-PROT entry is copyright. It is produced through a  
 collaboration between the Swiss Institute of Bioinformatics and  
 the EMBL outstation - the European Bioinformatics Institute.  
 The original entry is available from <http://www.expasy.ch/sprot>  
 and <http://www.ebi.ac.uk/sprot>  
 -----  
 [FUNCTION] Cell adhesion protein that participate in lymphocyte  
 recirculation by mediating the binding of lymphocytes to peripheral  
 lymph node vascular endothelial cells in an L-selectin-independent  
 fashion. Has a monoamine oxidase activity.  
 [CATALYTIC ACTIVITY]  $\text{RCH(2)NH(2)} + \text{H(2)O} + \text{O(2)} = \text{RCHO} + \text{NH(3)} + \text{H(2)O(2)}$ .  
 [COFACTOR] Binds 1 copper ion and 1 topaquinone per subunit.  
 [SUBUNIT] Homodimer.  
 [SUBCELLULAR LOCATION] Type II membrane protein.  
 [TISSUE SPECIFICITY] Most strongly expressed on the high  
 endothelial venules of peripheral lymph nodes and on hepatic  
 endothelia.  
 [PTM] Topaquinone (TPQ) is generated by copper-dependent  
 autoxidation of a specific tyrosyl residue (By similarity).  
 [PTM] N- and O-glycosylated.  
 [SIMILARITY] Belongs to the copper/topaquinone oxidase family.  
 FEATURES Location/Qualifiers  
     source 1..763  
           /organism="Homo sapiens"  
           /db\_xref="taxon:9606"  
     gene 1..763  
           /gene="AOC3"  
           /note="synonym: VAP1"  
     Protein 1..763

Region /gene="AOC3"  
/product="Membrane copper amine oxidase"  
/EC\_number="1.4.3.6"  
1..5  
/gene="AOC3"  
/region\_name="Domain"  
/note="Cytoplasmic (Potential)."  
Region 6..26  
/gene="AOC3"  
/region\_name="Transmembrane region"  
/note="Signal-anchor for type II membrane protein  
(Potential)."  
Region 27..763  
/gene="AOC3"  
/region\_name="Domain"  
/note="Extracellular (Potential)."  
Region 65..152  
/gene="AOC3"  
/region\_name="Copper amine oxidase, N2 domain"  
/note="Cu\_amine\_oxidN2"  
/db\_xref="CDD:3234"  
Site 137  
/gene="AOC3"  
/site\_type="glycosylation"  
/note="N-linked (GlcNAc...) (Potential)."  
Region 169..269  
/gene="AOC3"  
/region\_name="Copper amine oxidase, N3 domain"  
/note="Cu\_amine\_oxidN3"  
/db\_xref="CDD:6797"  
Site 232  
/gene="AOC3"  
/site\_type="glycosylation"  
/note="N-linked (GlcNAc...) (Potential)."  
Site 294  
/gene="AOC3"  
/site\_type="glycosylation"  
/note="N-linked (GlcNAc...) (Potential)."  
Region 308..724  
/gene="AOC3"  
/region\_name="Copper amine oxidase, enzyme domain"  
/note="Cu\_amine\_oxid"  
/db\_xref="CDD:25702"  
Region 317  
/gene="AOC3"  
/region\_name="Variant"  
/note="Y -> H (in dbSNP:438287). /FTId=VAR\_012064."  
Site 471  
/gene="AOC3"  
/site\_type="modified"  
/note="2',4',5'-topaquinone (By similarity)."  
Site 520  
/gene="AOC3"  
/site\_type="metal-binding"  
/note="Copper (Potential)."  
Site 522  
/gene="AOC3"  
/site\_type="metal-binding"  
/note="Copper (Potential)."  
Site 592

Site      /gene="AOC3"  
          /site\_type="glycosylation"  
          /note="N-linked (GlcNAc...) (Potential)."  
618  
Site      /gene="AOC3"  
          /site\_type="glycosylation"  
          /note="N-linked (GlcNAc...) (Potential)."  
666  
Site      /gene="AOC3"  
          /site\_type="glycosylation"  
          /note="N-linked (GlcNAc...) (Potential)."  
673  
Site      /gene="AOC3"  
          /site\_type="binding"  
          /note="Amiloride (By similarity)."  
684  
          /gene="AOC3"  
          /site\_type="metal-binding"  
          /note="Copper (Potential)."

## ORIGIN

```
1 mnqktilvll ilavitifal vcvllvgrgg dggepsqlph cpsvpsaqp wthpggsqf
61 adlsreelta vmrfltqrlg pglvd aaqar psdncv fsve lqlppkaaal ahldrgsppp
121 arealaivff grqpqp nvse lvvgplphps ymrdvt verh ggplpyhrrp vlfqeyldid
181 qmifnrelpq asgllhhccf ykhrgrnlvt mttaprglqs gdratwfgly ynissagffl
241 hhvglellvn hkaldparwt iqkvfyggry yds laqlea feaglvnvvl ipdngtggs
301 slkspvppgp applqfypqg prfsvqgsrv asslwtfsfg lgafsgprif dvrfggerlv
361 yeislqeala iyggns paam ttryvdggfg mgk yttpltr gvdcpylaty vdwhfllesq
421 apktirdafc vfeqnqglpl rrhhsdlysh yfgglaetvl vvrsmstlln ydyvwdtvfh
481 psgaieirfy atgyissaf l fgatgkygnq vsehtlgtvh thsahfkvdl dvaglenwvw
541 aedmvfvpma vpwspehqlq rlqvtrk lle meeqaaf lvg satprylyla snhsnkwghp
601 rgyriqmlsf ageplpq nss margfswery qlavtqrkee epssssvf nq ndpwaptvdf
661 sdfinnetia gkd lvawvta gflhiphaed ipntvtvgng vgfflrpy nf fdedpsfysa
721 dsiyfrgdqd agacevnpla clpqaaacap dlpafshggf shn
```

//

[Disclaimer](#) | [Write to the Help Desk](#)  
[NCBI](#) | [NLM](#) | [NIH](#)

Oct 29 2004 14:28:57

Database : SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp Vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	3405	63.6	763	6	Q9TTK6	Q9ttk6 bos taurus
2	3370	62.9	765	11	Q9R055	Q9r055 mus musculu
3	3174	59.2	1640	11	Q7TP38	Q7tp38 rattus norv
4	2390.5	44.6	757	11	Q812C9	Q812c9 mus musculu
5	2389.5	44.6	757	11	Q80WP3	Q80wp3 mus musculu
6	1404.5	26.2	751	11	Q8VC36	Q8vc36 mus musculu
7	1403.5	26.2	751	11	Q8JZQ5	Q8jzq5 mus musculu
8	1032	19.3	447	11	Q8R229	Q8r229 mus musculu
9	1029	19.2	218	5	Q94745	Q94745 schistosoma
10	992	18.5	820	11	Q8BZI2	Q8bzi2 mus musculu
11	771.5	14.4	821	3	Q86ZN4	Q86zn4 podospora a
12	721	13.5	218	5	Q25595	Q25595 clonorchis
13	707	13.2	218	5	Q9XYL9	Q9xyl9 fasciola gi
14	649	12.1	787	3	Q96X16	Q96x16 pichia past
15	615.5	11.5	271	11	Q920K6	Q920k6 cavia porce
16	518.5	9.7	218	6	Q9N0V4	Q9n0v4 bos taurus
17	514.5	9.6	219	13	Q7SZ23	Q7sz23 xenopus lae
18	513.5	9.6	219	13	Q90WM9	Q90wm9 xenopus lae
19	507.5	9.5	218	11	Q8R5I6	Q8r5i6 mus musculu
20	503.5	9.4	223	5	O97117	O97117 boophilus m
21	484.5	9.0	218	6	Q9TSM5	Q9tsm5 macaca fasc
22	481.5	9.0	218	6	Q9TSM4	Q9tsm4 macaca fasc
23	481.5	9.0	221	6	Q9BEB0	Q9beb0 macaca fusc
24	479.5	9.0	218	11	Q91Y83	Q91y83 cavia porce
25	479.5	9.0	218	11	Q8K0C3	Q8k0c3 mus musculu
26	472.5	8.8	204	11	Q9CW16	Q9cw16 mus musculu
27	470.5	8.8	219	5	Q8MWS0	Q8mws0 taenia soli
28	468.5	8.7	218	11	Q9WU21	Q9wu21 rattus norv
29	466.5	8.7	218	11	Q80W21	Q80w21 mus musculu

30	463.5	8.7	219	5	Q27653	Q27653 echinococcu
31	462.5	8.6	219	5	O16058	O16058 echinococcu
32	462	8.6	133	11	Q8CH13	Q8ch13 rattus norv
33	455.5	8.5	225	4	Q96HA3	Q96ha3 homo sapien
34	449.5	8.4	225	6	Q9BEA9	Q9bea9 macaca fusc
35	441.5	8.2	232	11	Q9D5J8	Q9d5j8 mus musculu
36	440.5	8.2	225	11	Q9Z1B2	Q9z1b2 rattus norv
37	439.5	8.2	195	4	Q05465	Q05465 homo sapien
38	417.5	7.8	219	5	Q8I9R9	Q8i9r9 sarcoptes s
39	415.5	7.8	188	6	Q9MZB4	Q9mzb4 capra hircu
40	388	7.2	741	10	Q9C6V7	Q9c6v7 arabidopsis
41	386.5	7.2	219	5	Q9U582	Q9u582 psoroptes o
42	381	7.1	654	16	Q989X2	Q989x2 rhizobium l
43	372.5	7.0	668	10	O48552	O48552 arabidopsis
44	371.5	6.9	557	10	Q8GZ62	Q8gz62 arabidopsis
45	368.5	6.9	660	17	Q97XM1	Q97xm1 sulfolobus

# ALIGNMENTS

## RESULT 1

### Q9TTK6

ID Q9TTK6 PRELIMINARY; PRT; 763 AA.  
AC Q9TTK6;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Semicarbazide-sensitive amine oxidase (EC 1.4.3.6) (Copper amine  
DE oxidase).  
GN SSAO.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Iwabuki H., Matsumura K., Mure M., Kuroda S., Tanizawa K.;  
RT "Molecular cloning of semicarbazide-sensitive amine oxidase gene from  
RT Bovine aorta."  
RL Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.  
CC -!- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +  
CC H(2)O(2).  
CC -!- COFACTOR: BINDS 1 COPPER ION AND 1 TOPAQUINONE PER SUBUNIT (BY  
CC SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE COPPER/TOPAQUINONE OXIDASE FAMILY.  
DR EMBL; AB019242; BAA88896.1; -.  
DR GO; GO:0008122; F:amine oxidase (copper-containing) activity; IEA.  
DR GO; GO:0005507; F:copper ion binding; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR InterPro; IPR000269; CuNH\_oxidase.  
DR Pfam; PF01179; Cu\_amine\_oxid; 1.  
DR Pfam; PF02727; Cu\_amine\_oxidN2; 1.  
DR Pfam; PF02728; Cu\_amine\_oxidN3; 1.  
DR PRINTS; PR00766; CUDAOXIDASE.  
DR PROSITE; PS01164; COPPER\_AMINE\_OXID\_1; 1.

DR PROSITE; PS01165; COPPER\_AMINE\_OXID\_2; 1.  
KW Copper; Oxidoreductase; TPQ.  
SQ SEQUENCE 763 AA; 84500 MW; DB61ED9A89E71E90 CRC64;

Query Match 63.6%; Score 3405; DB 6; Length 763;  
Best Local Similarity 85.1%; Pred. No. 1.2e-267;  
Matches 626; Conservative 44; Mismatches 66; Indels 0; Gaps 0;

Qy 263 QGGDGGEPSQLPHCPSPSPAQPWTHPGQSQLFADLSREELTAVMRFLTQRLGPGLVDAA 322  
:||||| :||| :|| | ||||| :||||| ||||| :||| |||||  
Db 28 RGGDGGEASQPHYCPSGTPSVQPWTHPGQNQLFADLSREELTAVMSFLTQKLGPDLDVDA 87

Qy 323 QARPSDNCVFSVELQLPPKAAALAHLDRGSPPPAREALAIVFFGRQPQPNVSELVVGPLP 382  
||||| :||||| ||||| ||||| :||||| ||||| :||||| |||||  
Db 88 QARPSDNCIFSVELQLPPKAAALAHLDRRSPPPAREALAIVFFGGQPQPNVTELVVGPLP 147

Qy 383 HPSYMRDVTVERHGGPLPYHRRPVLFQEYLDIDQMIFNRELPAASGLLHHCCFYKHRGRN 442  
||||| :||||| :||||| ||||| :||| :||| ||| |  
Db 148 QPSYMRDVTVERHGGPLPYRRPVLLREYLDIDQMIFNRELPAAGVLHHCCSYKQGGGN 207

Qy 443 LVTMTTAPRGLQSGDRATWFGLYYNISGAGFFLHHVGLLELVNHHKALDPAWRTIQKVFIQ 502  
||||| :||||| ||||| :||| ||||| :||||| :||||| :|||  
Db 208 LVTMTTAPRGLQSGDRATWFGLYYNISGAGYYLHPVGLLELVNHHKALDPAQWTIQKVFIQ 267

Qy 503 GRYYDSLAQLEAQFEAGLVNVVLIIPDNGTGGSSWSLKSPVPPGPAPPLQFYPPQPRFSVQG 562  
||| :||||| ||||| ||||| :||| :||||| ||||| ||||| :||| |||||  
Db 268 GRYYESLAQLEEQFEAGRVNVVVIIPNNGTGGSSWSLKSPVPPGPPTPLQFHPQGRFSVQG 327

Qy 563 SRVASSLWTFSGFLGAFSGPRIFDVRFQGERLVYEISLQEALAIYGGNSPAAMTTRYVDG 622  
||| ||||| :||||| :||||| ||||| :||||| :||||| :||| ||| :|||  
Db 328 SRVTSSLWTFSGFLGAFSGPRIFDIRFQGERLAYEISLQEAVALIYGGNTPAAMLTRYMDG 387

Qy 623 GFGMGKYTTPLTRGVDCPYLATYVDWHFLLSQAPKTIRDAFCVFEQNQGLPLRRHHSDDL 682  
||||| : ||||| ||||| ||||| ||||| :||| : ||||| : ||||| |||||  
Db 388 CFGMGKFATPLTRGVDCPYLATYVDWHFLLSQAPRTLHDAFCVFEQNKGLPLRRHHSDF 447

Qy 683 YSHYFGGLAETVLVVRSMSTLLNYDYVWDTVFHPSGAIEIRFYATGYISSAFLFGATGKY 742  
| ||||| : ||||| ||| : ||||| ||||| ||||| : ||||| : ||||| ||| |||  
Db 448 ISQYFGGVVETVLVFRSVSTLLNYDYVWDMVFHPNGAIEVKFHATGYISSAFFFGTAQKY 507

Qy 743 GNQVSEHTLGTVHTHSAHFVLDLDVAGLENVWVAEDMVFPMAVPWSPQHQQLQVTRK 802  
||| | : ||||| ||||| : ||||| ||||| ||||| ||| ||||| : ||||| |||  
Db 508 GNQVRENTLGTVHTHSAHYKVDLDVGLENVWVAEDMAFVPTTPWSPQHQQLQVTRK 567

Qy 803 LLEMEEQAFLVGSATPRYLYLASNHSNKWGHPRGYRIQMLSFAGEPLPQNSSMARGFSW 862  
|| ||||| : | : ||||| ||||| ||||| : ||| ||||| ||| |||  
Db 568 QLETEEQAFLVGSATPRYLYLASNHSNKWGHPRGYRIQTVSFAGRPLPQNSSSTERAISW 627

Qy 863 ERYQLAVTQRKEEEPSSSSVFVNQNDPWAPTVDVDFINNETIAGKDLVAWVTAGFLHIPH 922  
||||| ||||| ||||| ||||| ||||| ||||| : ||||| ||||| ||||| ||||| |||||  
Db 628 GRYQLAVTQRKETEPSSSSVFVNQNDPWPTVDFADFINNETIAGKDLVAWVTAGFLHIPH 687

Qy 923 AEDIPNTVTVGNGVGFFLRPNFFDEDPFSYADSIFRGDQDAGACEVNPLACLPQAAA 982  
||||| ||||| ||||| ||||| ||||| ||||| : ||||| : ||||| ||||| : |||  
Db 688 AEDIPNTVTVGNGVGFFLRPNFFDEDPFSINSADSIYFQKHQDAGSCEVNSLACLPKDPA 747

Qy 983 CAPDLPAFSGHGGFSHN 998

Db

|||||  
748 CAPDLPAFSHGGFTN 763



Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result		%					
No.	Score	Query Match	Length	DB	ID	Description	
1	3967	74.1	763	1	AOC3_HUMAN	Q16853	homo sapien
2	3365	62.8	765	1	AOC3_MOUSE	O70423	mus musculu
3	3338	62.3	762	1	AOCY_BOVIN	O46406	bos taurus
4	3326	62.1	762	1	AOCX_BOVIN	Q29437	bos taurus
5	2556	47.7	756	1	AOC2_HUMAN	O75106	homo sapien
6	1398.5	26.1	746	1	ABP_RAT	P36633	rattus norv
7	1394	26.0	751	1	ABP_HUMAN	P19801	homo sapien
8	1315	24.5	321	1	AOC3_RAT	O08590	rattus norv
9	1144	21.4	218	1	GT26_SCHJA	P08515	schistosoma
10	985	18.4	218	1	GT27_SCHMA	P35661	schistosoma
11	922	17.2	218	1	GT26_SCHMA	P15964	schistosoma
12	719.5	13.4	220	1	GT29_FASHE	P56598	fasciola he
13	708	13.2	217	1	GT27_FASHE	P31670	fasciola he
14	690	12.9	217	1	GT28_FASHE	P31671	fasciola he
15	689	12.9	217	1	GT26_FASHE	P30112	fasciola he
16	489.5	9.1	217	1	GTM1_MOUSE	P10649	mus musculu
17	489.5	9.1	217	1	GTM1_RAT	P04905	rattus norv
18	485.5	9.1	217	1	GTMU_RABIT	P46409	oryctolagus
19	485.5	9.1	218	1	GTM6_MOUSE	O35660	mus musculu
20	484.5	9.0	217	1	GTM2_HUMAN	P28161	homo sapien
21	483.5	9.0	218	1	GTM4_HUMAN	Q03013	homo sapien
22	481.5	9.0	217	1	GTMU_CAVPO	P16413	cavia porce
23	476.5	8.9	217	1	GTMU_CRILO	Q00285	cricetulus
24	476.5	8.9	217	1	GTMU_MESAU	P30116	mesocricetu
25	475.5	8.9	217	1	GTM2_MOUSE	P15626	mus musculu
26	473.5	8.8	217	1	GTM1_HUMAN	P09488	homo sapien

Database : PIR\_78:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		Length	DB	ID	Description
		Match					
1	3967	74.1		763	2	JC5234	amine oxidase (cop
2	3326	62.1		762	2	A54411	amine oxidase (cop
3	1398.5	26.1		746	2	S34656	amine oxidase (cop
4	1394	26.0		751	2	A54053	amine oxidase (cop
5	1140	21.3		219	2	A26484	glutathione transf
6	985	18.4		218	2	A45556	glutathione S-tran
7	922	17.2		218	2	A45523	glutathione transf
8	663	12.4		209	2	A48388	glutathione S-tran
9	489.5	9.1		218	2	S33860	glutathione transf
10	489.5	9.1		218	2	A29794	glutathione transf
11	486.5	9.1		218	2	S32425	glutathione transf
12	485.5	9.1		218	2	S65674	glutathione transf
13	484.5	9.0		218	2	A39375	glutathione transf
14	483.5	9.0		218	2	A47486	glutathione transf
15	481.5	9.0		217	2	JX0095	glutathione transf
16	476.5	8.9		218	2	S13202	glutathione transf
17	476.5	8.9		218	2	A23732	glutathione transf
18	475.5	8.9		218	2	B34159	glutathione transf
19	473.5	8.8		218	2	S01719	glutathione transf
20	471.5	8.8		218	2	A46048	glutathione transf

Database :       Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result		%					
No.	Score	Query	Match	Length	ID		Description
1	5357	100.0	998	13	US-10-081-408-20		Sequence 20, Appl
2	3967	74.1	763	13	US-10-081-408-2		Sequence 2, Appli
3	2480.5	46.3	729	9	US-09-919-497-51		Sequence 51, Appl
4	1368.5	25.5	770	16	US-10-408-765A-125		Sequence 125, App
5	1184	22.1	504	14	US-10-050-902-289		Sequence 289, App
6	1184	22.1	504	14	US-10-050-898-289		Sequence 289, App
7	1179.5	22.0	419	14	US-10-050-902-318		Sequence 318, App
8	1179.5	22.0	419	14	US-10-050-898-318		Sequence 318, App
9	1176	22.0	692	15	US-10-318-308-1		Sequence 1, Appli
10	1171.5	21.9	514	12	US-10-325-810-605		Sequence 605, App

Database : Issued\_Patents\_AA:\*  
 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	% Query		Length	DB	ID	Description
	Score	Match				
1	1180	22.0	447	4	US-08-506-296B-73	Sequence 73, Appl
2	1175.5	21.9	644	4	US-08-506-296B-65	Sequence 65, Appl
3	1175	21.9	536	3	US-08-974-180-20	Sequence 20, Appl
4	1171.5	21.9	514	3	US-08-974-549A-605	Sequence 605, App
5	1171.5	21.9	514	4	US-08-912-951-319	Sequence 319, App
6	1171.5	21.9	514	4	US-09-402-181B-605	Sequence 605, App
7	1171.5	21.9	514	4	US-09-721-456-605	Sequence 605, App
8	1168	21.8	1252	4	US-10-012-762-20	Sequence 20, Appl
9	1168	21.8	1252	4	US-09-704-036B-20	Sequence 20, Appl
10	1167.5	21.8	362	1	US-08-395-507-2	Sequence 2, Appli
11	1165	21.7	517	3	US-08-974-549A-606	Sequence 606, App
12	1165	21.7	517	4	US-08-912-951-320	Sequence 320, App
13	1165	21.7	517	4	US-09-402-181B-606	Sequence 606, App
14	1165	21.7	517	4	US-09-721-456-606	Sequence 606, App
15	1164.5	21.7	443	4	US-08-506-296B-76	Sequence 76, Appl

Database : A\_Geneseq\_29Jan04:\*  
 1: geneseqp1980s:\*  
 2: geneseqp1990s:\*  
 3: geneseqp2000s:\*  
 4: geneseqp2001s:\*  
 5: geneseqp2002s:\*  
 6: geneseqp2003as:\*  
 7: geneseqp2003bs:\*  
 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	5357	100.0	998	5	AAE26690	Aae26690 S. japoni
2	3967	74.1	763	2	AAY03219	Aay03219 Amino aci
3	3967	74.1	763	5	AAE26686	Aae26686 Human sem
4	2480.5	46.3	729	5	AAU84261	Aau84261 Human end
5	1398.5	26.1	746	7	ADE64055	Ade64055 Rat Prote
6	1394	26.0	751	7	ADE64057	Ade64057 Human Pro
7	1184	22.1	504	5	ABG94324	Abg94324 Lymphotox
8	1184	22.1	504	5	ABG80636	Abg80636 Mouse lym
9	1182.5	22.1	697	7	ADE52287	Ade52287 Human GST
10	1180	22.0	447	5	AAU10670	Aau10670 Mouse L1C
11	1179.5	22.0	419	5	ABG94337	Abg94337 GST-PS-C-
12	1179.5	22.0	419	5	ABG80649	Abg80649 Mouse RAN
13	1176.5	22.0	843	7	ADC51505	Adc51505 Human GST

RESULT 2

AA03219

ID AA03219 standard; protein; 763 AA.

XX

AC AA03219;

XX

DT 21-JUN-1999 (first entry)

XX

DE Amino acid sequence of the vascular adhesion protein-1.

XX

KW Human; vascular adhesion protein-1; VAP-1; endothelial cell; lymphocyte;  
 KW inhibition; amine oxidase.

XX

OS Homo sapiens.

XX

Key	Location/Qualifiers
FT Peptide	1. .25
FT	/note= "N-terminal tryptic peptide"
FT Domain	5. .27
FT	/note= "transmembrane domain"
FT Misc-difference	43
FT	/note= "potential O-glycosylation site"
FT Misc-difference	47
FT	/note= "potential O-glycosylation site"
FT Peptide	74. .78
FT	/note= "V8 peptide"
FT Misc-difference	137
FT	/note= "potential N-glycosylation site"
FT Peptide	170. .174
FT	/note= "V8 peptide"
FT Misc-difference	232
FT	/note= "potential N-glycosylation site"
FT Peptide	259. .262
FT	/note= "V8 peptide"
FT Peptide	264. .277
FT	/note= "V8 peptide"
FT Misc-difference	294
FT	/note= "potential N-glycosylation site"
FT Peptide	323. .329
FT	/note= "V8 peptide"
FT Peptide	359. .361
FT	/note= "V8 peptide"
FT Peptide	368. .375
FT	/note= "V8 peptide"
FT Peptide	384. .393
FT	/note= "V8 peptide"
FT Peptide	562. .566
FT	/note= "V8 peptide"
FT Misc-difference	592
FT	/note= "potential N-glycosylation site"
FT Peptide	597. .601
FT	/note= "V8 peptide"
FT Misc-difference	618
FT	/note= "potential N-glycosylation site"
FT Peptide	624. .637
FT	/note= "V8 peptide"
FT Misc-difference	666

```

FT      /note= "potential N-glycosylation site"
FT      Misc-difference 679
FT      /note= "potential O-glycosylation site"
XX
PN      WO9853049-A1.
XX
PD      26-NOV-1998.
XX
PF      22-MAY-1998;    98WO-FI000429.
XX
PR      23-MAY-1997;    97US-00862433.
XX
PA      (BIOT-) BIOTIE THERAPIES LTD.
XX
PI      Jalkanen S,  Salmi M,  Smith DJ,  Bono P;
XX
DR      WPI; 1999-131690/11.
DR      N-PSDB; AAX28632.
XX
PT      Nucleic acid encoding a novel human endothelial cell adhesion protein
PT      designated VAP-1 - having an adhesive function and an amine oxidase
PT      function useful for manipulating VAP-1 mediated binding of endothelial
PT      cells to lymphocytes.
XX
PS      Claim 1; Fig 1; 66pp; English.
XX
CC      This is the nucleotide sequence encoding the human vascular adhesion
CC      protein-1 (VAP-1) used in the method of the invention. The method
CC      involves manipulating VAP-1 mediated binding of endothelial cells to
CC      lymphocytes which comprises inhibiting the enzymatic activity of amine
CC      oxidase in endothelial cells, and potentiating the enzymatic activity of
CC      endothelial cells
XX
SQ      Sequence 763 AA;

Query Match          74.1%;  Score 3967;  DB 2;  Length 763;
Best Local Similarity 99.9%;  Pred. No. 0;
Matches 735;  Conservative 1;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      263 QGGDGGEP SQLPHCPSVSPSAQPWTHPGQSQLFADLSREELTAVMRFLTQRLGPGLVDA 322
      :|||||
Db      28  RGGDGGEP SQLPHCPSVSPSAQPWTHPGQSQLFADLSREELTAVMRFLTQRLGPGLVDA 87

Qy      323 QARPSDNCVFSVELQLPPKAAALHLDRGSPPPAREALAI VFFGRQPQPNVSELVVGPLP 382
      ||||||
Db      88  QARPSDNCVFSVELQLPPKAAALHLDRGSPPPAREALAI VFFGRQPQPNVSELVVGPLP 147

Qy      383 HPSYMRDVTVERHGGPLPYHRRPVLFQEYLDIDQMIFNREL PQASGLLHHCCFYKHRGRN 442
      ||||||
Db      148 HPSYMRDVTVERHGGPLPYHRRPVLFQEYLDIDQMIFNREL PQASGLLHHCCFYKHRGRN 207

Qy      443 LVTMTTAPRGLQSGDRATWFGLYYNISGAGFFLHHVGLELLVNHKALDPARWTIQKVFYQ 502
      ||||||
Db      208 LVTMTTAPRGLQSGDRATWFGLYYNISGAGFFLHHVGLELLVNHKALDPARWTIQKVFYQ 267

Qy      503 GRYYDSLAEQLEAQFEAGLVNVVLIPDNGTGGSWSLKSPVPPGPAPPLQFYPPQGRFSVQG 562
      ||||||

```

Db	268	GRYYDSLAQLEAQFEAGLVNVVLI	PDNGTGGSWSLKSPVPPGPAPPLQFYPQGP	PRFSVQG	327
Qy	563	SRVASSLWTF	FSFGLGAFSGPRIFDVR	FQGERLVYEISLQEALAIYGGNSPAAM	TRYVDG 622
Db	328	SRVASSLWTF	FSFGLGAFSGPRIFDVR	FQGERLVYEISLQEALAIYGGNSPAAM	TRYVDG 387
Qy	623	GFGMGKYTTPL	TRGVDCPYLATYVDWHF	LLESQAPKTIRDAFCVFEQNQGLPLRR	HSDL 682
Db	388	GFGMGKYTTPL	TRGVDCPYLATYVDWHF	LLESQAPKTIRDAFCVFEQNQGLPLRR	HSDL 447
Qy	683	YSHYFGGLAET	VLVVRSMSTLLNYDYVWD	TVFHPSGAIEIRFYATGYISSAFL	FGATGKY 742
Db	448	YSHYFGGLAET	VLVVRSMSTLLNYDYVWD	TVFHPSGAIEIRFYATGYISSAFL	FGATGKY 507
Qy	743	GNQVSEHTLGT	VHTHSAHFKVDLDVAG	LENWVWAEDMVFPMAVPWSPEHQLQ	RLQVTRK 802
Db	508	GNQVSEHTLGT	VHTHSAHFKVDLDVAG	LENWVWAEDMVFPMAVPWSPEHQLQ	RLQVTRK 567
Qy	803	LLEMEEQAAFL	VGSATPRYLYLASNHS	NKWGHPRGYRIQMLSFAGEPLPQN	SSMARGFSW 862
Db	568	LLEMEEQAAFL	VGSATPRYLYLASNHS	NKWGHPRGYRIQMLSFAGEPLPQN	SSMARGFSW 627
Qy	863	ERYQLAVTQR	KEEEPSSSSVFNQNDP	WAPTVD	FSDFINNETIAGKDLVAWVTAGFLHIPH 922
Db	628	ERYQLAVTQR	KEEEPSSSSVFNQNDP	WAPTVD	FSDFINNETIAGKDLVAWVTAGFLHIPH 687
Qy	923	AEDIPNTVT	VGNGVGFFLRPYNFF	DEDPSFY	SADSIYFRGDQDAGACEVNPLACLPQAAA 982
Db	688	AEDIPNTVT	VGNGVGFFLRPYNFF	DEDPSFY	SADSIYFRGDQDAGACEVNPLACLPQAAA 747
Qy	983	CAPDLPAF	SHGGFSHN		998
Db	748	CAPDLPAF	SHGGFSHN		763



RESULT 1

AOC3\_HUMAN

ID AOC3\_HUMAN STANDARD; PRT; 763 AA.  
AC Q16853;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Membrane copper amine oxidase (EC 1.4.3.6) (Vascular adhesion protein-  
DE 1) (VAP-1) (HPAO).  
GN AOC3 OR VAP1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=97128319; PubMed=8972912;  
RA Zhang X., McIntire W.S.;  
RT "Cloning and sequencing of a copper-containing, topaquinone-  
RT containing monoamine oxidase from human placenta.";  
RL Gene 179:279-286(1996).  
RN [2]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Lung;  
RX MEDLINE=98317014; PubMed=9653080;  
RA Smith D.J., Salmi M., Bono P., Hellman J., Leu T., Jalkanen S.;  
RT "Cloning of vascular adhesion protein 1 reveals a novel  
RT multifunctional adhesion molecule.";  
RL J. Exp. Med. 188:17-27(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PNS;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- FUNCTION: Cell adhesion protein that participate in lymphocyte  
CC recirculation by mediating the binding of lymphocytes to  
CC peripheral lymph node vascular endothelial cells in an L-selectin-

CC independent fashion. Has a monoamine oxidase activity.  
 CC -!- CATALYTIC ACTIVITY:  $\text{RCH(2)NH(2)} + \text{H(2)O} + \text{O(2)} = \text{RCHO} + \text{NH(3)} + \text{H(2)O(2)}$ .  
 CC -!- COFACTOR: Binds 1 copper ion and 1 topaquinone per subunit.  
 CC -!- SUBUNIT: Homodimer.  
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein.  
 CC -!- TISSUE SPECIFICITY: MOST STRONGLY EXPRESSED ON THE HIGH  
 CC ENDOTHELIAL VENULES OF PERIPHERAL LYMPH NODES AND ON HEPATIC  
 CC ENDOTHELIA.  
 CC -!- PTM: Topaquinone (TPQ) is generated by copper-dependent  
 CC autoxidation of a specific tyrosyl residue (By similarity).  
 CC -!- PTM: N- and O-glycosylated.  
 CC -!- SIMILARITY: Belongs to the copper/topaquinone oxidase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U39447; AAC50919.1; -.  
 DR EMBL; AF067406; AAC25170.1; -.  
 DR EMBL; BC050549; AAH50549.1; -.  
 DR PIR; JC5234; JC5234.  
 DR Genew; HGNC:550; AOC3.  
 DR MIM; 603735; -.  
 DR GO; GO:0016021; C:integral to membrane; TAS.  
 DR GO; GO:0005886; C:plasma membrane; TAS.  
 DR GO; GO:0008122; F:amine oxidase (copper-containing) activity; TAS.  
 DR GO; GO:0005489; F:electron transporter activity; TAS.  
 DR GO; GO:0009308; P:amine metabolism; TAS.  
 DR InterPro; IPR000269; CuNH\_oxidase.  
 DR Pfam; PF01179; Cu\_amine\_oxid; 1.  
 DR Pfam; PF02727; Cu\_amine\_oxidN2; 1.  
 DR Pfam; PF02728; Cu\_amine\_oxidN3; 1.  
 DR PRINTS; PR00766; CUDAOXIDASE.  
 DR PROSITE; PS01164; COPPER\_AMINE\_OXID\_1; 1.  
 DR PROSITE; PS01165; COPPER\_AMINE\_OXID\_2; 1.  
 KW Oxidoreductase; Copper; TPQ; Glycoprotein; Transmembrane;  
 KW Signal-anchor; Cell adhesion; Polymorphism; Metal-binding.  
 FT DOMAIN 1 5 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 6 26 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT (POTENTIAL).  
 FT DOMAIN 27 763 EXTRACELLULAR (POTENTIAL).  
 FT MOD\_RES 471 471 TOPAQUINONE (BY SIMILARITY).  
 FT METAL 520 520 COPPER (POTENTIAL).  
 FT METAL 522 522 COPPER (POTENTIAL).  
 FT METAL 684 684 COPPER (POTENTIAL).  
 FT BINDING 673 673 AMILORIDE (BY SIMILARITY).  
 FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 592 592 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 666 666 N-LINKED (GLCNAC. . .) (POTENTIAL).

Seq ID 10:20

FT VARIANT 317 317 Y -> H (in dbSNP:438287).  
FT /FTId=VAR\_012064.  
SQ SEQUENCE 763 AA; 84621 MW; 58AD55605EC9D228 CRC64;

Query Match 74.1%; Score 3967; DB 1; Length 763;  
Best Local Similarity 99.9%; Pred. No. 4.7e-286;  
Matches 735; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy 263 QGGDGGEPSQLPHCPSPVSPSAQPWTHPGSQLFADLSREELTAVMRFLTQRLGPGLVDAA 322
      :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28 RGGDGGEPSQLPHCPSPVSPSAQPWTHPGSQLFADLSREELTAVMRFLTQRLGPGLVDAA 87

Qy 323 QARPSDNCVFSVELQLPPKAAALAHLDGRGSPPPAREALAIVFFGRQPQPNVSELVVGPLP 382
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 88 QARPSDNCVFSVELQLPPKAAALAHLDGRGSPPPAREALAIVFFGRQPQPNVSELVVGPLP 147

Qy 383 HPSYMRDVTVERHGGPLPYHRRPVLFQEYLDIDQMIFNREL PQASGLLHHCCFYKHRGRN 442
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 148 HPSYMRDVTVERHGGPLPYHRRPVLFQEYLDIDQMIFNREL PQASGLLHHCCFYKHRGRN 207

Qy 443 LVTMTTAPRGLQSGDRATWFGLYYNISGAGFFLHHVGLELLVNHKALDPARWTIQKFYQ 502
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 208 LVTMTTAPRGLQSGDRATWFGLYYNISGAGFFLHHVGLELLVNHKALDPARWTIQKFYQ 267

Qy 503 GRYYDSLAEQLEAQFEAGLVNVVLIPDNGTGGSWSLKSPVPPGPAPPLQFYPQGPRFSVQG 562
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 268 GRYYDSLAEQLEAQFEAGLVNVVLIPDNGTGGSWSLKSPVPPGPAPPLQFYPQGPRFSVQG 327

Qy 563 SRVASSLWTFSGFLGAFSGPRIFDVRFQGERLVYEISLQEALAIYGGNSPAAMTTRYVDG 622
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 328 SRVASSLWTFSGFLGAFSGPRIFDVRFQGERLVYEISLQEALAIYGGNSPAAMTTRYVDG 387

Qy 623 GFGMGKYTTPLTRGVDCPYLATYVDWHFLLESQAPKTIRDAFCVFEQNQGLPLRRHSD 682
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 388 GFGMGKYTTPLTRGVDCPYLATYVDWHFLLESQAPKTIRDAFCVFEQNQGLPLRRHSD 447

Qy 683 YSHYFGGLAETVLVVRSMSTLLNYDYVWDTVFHPSGAIEIRFYATGYISSAFLEFGATGKY 742
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 448 YSHYFGGLAETVLVVRSMSTLLNYDYVWDTVFHPSGAIEIRFYATGYISSAFLEFGATGKY 507

Qy 743 GNQVSEHTLGTVHTHSAHFKVDLDVAGLENWVWAEDMVFPMAVPWSPEHQLQRLQVTRK 802
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 508 GNQVSEHTLGTVHTHSAHFKVDLDVAGLENWVWAEDMVFPMAVPWSPEHQLQRLQVTRK 567

Qy 803 LLEMEEQAAFLVGSATPRYLYLASNHSNKWGHPRGYRIQMLSFAGEPLPQNSSMARGFSW 862
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 568 LLEMEEQAAFLVGSATPRYLYLASNHSNKWGHPRGYRIQMLSFAGEPLPQNSSMARGFSW 627

Qy 863 ERYQLAVTQRKEEEPSSSVFNQNDPWAPTVDVDFINNETIAGKDLVAWVTAGFLHIPH 922
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 628 ERYQLAVTQRKEEEPSSSVFNQNDPWAPTVDVDFINNETIAGKDLVAWVTAGFLHIPH 687

Qy 923 AEDIPNTVTVGNGVGFFLRPNYFFDEDPFYSADSIYFRGDQDAGACEVNPLACLPQAAA 982
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 688 AEDIPNTVTVGNGVGFFLRPNYFFDEDPFYSADSIYFRGDQDAGACEVNPLACLPQAAA 747

Qy 983 CAPDLPAFSGGFSHN 998
```

Db

|||||  
748 CAPDLPAFSHGGFSHN 763

&lt;220&gt;

&lt;223&gt; Recombinant construct

&lt;400&gt; 20

Met Asp Trp Leu Arg Asn Leu Leu Phe Leu Met Ala Ala Ala Gln Ser  
 1 5 10 15  
 Ile Asn Ala Ala Gln His Asp Glu Ala Val Asp Asn Lys Phe Asn Lys  
 20 25 30  
 Glu Gln Gln Asn Ala Ser Pro Ile Leu Gly Tyr Trp Lys Ile Lys Gly  
 35 40 45  
 Leu Val Gln Pro Thr Arg Leu Leu Glu Tyr Leu Glu Glu Lys Tyr  
 50 55 60  
 Glu Glu His Leu Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys  
 65 70 75 80  
 Lys Phe Glu Leu Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp  
 85 90 95  
 Gly Asp Val Lys Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala  
 100 105 110  
 Asp Lys His Asn Met Leu Gly Gly Ser Pro Lys Glu Arg Ala Glu Ile  
 115 120 125  
 Ser Met Leu Glu Gly Ala Val Leu Asp Ile Arg Tyr Gly Val Ser Arg  
 130 135 140  
 Ile Ala Tyr Ser Lys Asp Phe Glu Thr Leu Lys Val Asp Phe Leu Ser  
 145 150 155 160  
 Lys Leu Pro Glu Met Leu Lys Met Phe Glu Asp Arg Leu Ser His Lys  
 165 170 175  
 Thr Tyr Leu Asn Gly Asp His Val Thr His Pro Asp Phe Met Leu Tyr  
 180 185 190  
 Asp Ala Leu Asp Val Val Leu Tyr Met Asp Pro Met Cys Leu Asp Ala  
 195 200 205  
 Phe Pro Lys Leu Val Ser Phe Lys Lys Arg Ile Glu Ala Ile Pro Gln  
 210 215 220  
 Ile Asp Lys Tyr Leu Lys Ser Ser Lys Tyr Ile Ala Trp Pro Leu Gln  
 225 230 235 240  
 Gly Trp Gln Ala Thr Phe Gly Gly Gly Asp His Pro Pro Lys Ser Gln  
 245 250 255  
 Ser Gln Glu Ala Leu Phe Gln Gly Gly Asp Gly Gly Glu Pro Ser Gln  
 260 265 270  
 Leu Pro His Cys Pro Ser Val Ser Pro Ser Ala Gln Pro Trp Thr His  
 275 280 285  
 Pro Gly Gln Ser Gln Leu Phe Ala Asp Leu Ser Arg Glu Glu Leu Thr  
 290 295 300  
 Ala Val Met Arg Phe Leu Thr Gln Arg Leu Gly Pro Gly Leu Val Asp  
 305 310 315 320  
 Ala Ala Gln Ala Arg Pro Ser Asp Asn Cys Val Phe Ser Val Glu Leu  
 325 330 335  
 Gln Leu Pro Pro Lys Ala Ala Ala Leu Ala His Leu Asp Arg Gly Ser  
 340 345 350  
 Pro Pro Pro Ala Arg Glu Ala Leu Ala Ile Val Phe Phe Gly Arg Gln  
 355 360 365  
 Pro Gln Pro Asn Val Ser Glu Leu Val Val Gly Pro Leu Pro His Pro  
 370 375 380  
 Ser Tyr Met Arg Asp Val Thr Val Glu Arg His Gly Gly Pro Leu Pro  
 385 390 395 400  
 Tyr His Arg Arg Pro Val Leu Phe Gln Glu Tyr Leu Asp Ile Asp Gln  
 405 410 415

signal

signal peptide

fusion partner

quar

SC  
cleavage

SSAD

Met Ile Phe Asn Arg Glu Leu Pro Gln Ala Ser Gly Leu Leu His His  
 420 425 430  
 Cys Cys Phe Tyr Lys His Arg Gly Arg Asn Leu Val Thr Met Thr Thr  
 435 440 445  
 Ala Pro Arg Gly Leu Gln Ser Gly Asp Arg Ala Thr Trp Phe Gly Leu  
 450 455 460  
 Tyr Tyr Asn Ile Ser Gly Ala Gly Phe Phe Leu His His Val Gly Leu  
 465 470 475 480  
 Glu Leu Leu Val Asn His Lys Ala Leu Asp Pro Ala Arg Trp Thr Ile  
 485 490 495  
 Gln Lys Val Phe Tyr Gln Gly Arg Tyr Tyr Asp Ser Leu Ala Gln Leu  
 500 505 510  
 Glu Ala Gln Phe Glu Ala Gly Leu Val Asn Val Val Leu Ile Pro Asp  
 515 520 525  
 Asn Gly Thr Gly Gly Ser Trp Ser Leu Lys Ser Pro Val Pro Pro Gly  
 530 535 540  
 Pro Ala Pro Pro Leu Gln Phe Tyr Pro Gln Gly Pro Arg Phe Ser Val  
 545 550 555 560  
 Gln Gly Ser Arg Val Ala Ser Ser Leu Trp Thr Phe Ser Phe Gly Leu  
 565 570 575  
 Gly Ala Phe Ser Gly Pro Arg Ile Phe Asp Val Arg Phe Gln Gly Glu  
 580 585 590  
 Arg Leu Val Tyr Glu Ile Ser Leu Gln Glu Ala Leu Ala Ile Tyr Gly  
 595 600 605  
 Gly Asn Ser Pro Ala Ala Met Thr Thr Arg Tyr Val Asp Gly Gly Phe  
 610 615 620  
 Gly Met Gly Lys Tyr Thr Thr Pro Leu Thr Arg Gly Val Asp Cys Pro  
 625 630 635 640  
 Tyr Leu Ala Thr Tyr Val Asp Trp His Phe Leu Leu Glu Ser Gln Ala  
 645 650 655  
 Pro Lys Thr Ile Arg Asp Ala Phe Cys Val Phe Glu Gln Asn Gln Gly  
 660 665 670  
 Leu Pro Leu Arg Arg His His Ser Asp Leu Tyr Ser His Tyr Phe Gly  
 675 680 685  
 Gly Leu Ala Glu Thr Val Leu Val Val Arg Ser Met Ser Thr Leu Leu  
 690 695 700  
 Asn Tyr Asp Tyr Val Trp Asp Thr Val Phe His Pro Ser Gly Ala Ile  
 705 710 715 720  
 Glu Ile Arg Phe Tyr Ala Thr Gly Tyr Ile Ser Ser Ala Phe Leu Phe  
 725 730 735  
 Gly Ala Thr Gly Lys Tyr Gly Asn Gln Val Ser Glu His Thr Leu Gly  
 740 745 750  
 Thr Val His Thr His Ser Ala His Phe Lys Val Asp Leu Asp Val Ala  
 755 760 765  
 Gly Leu Glu Asn Trp Val Trp Ala Glu Asp Met Val Phe Val Pro Met  
 770 775 780  
 Ala Val Pro Trp Ser Pro Glu His Gln Leu Gln Arg Leu Gln Val Thr  
 785 790 795 800  
 Arg Lys Leu Leu Glu Met Glu Glu Gln Ala Ala Phe Leu Val Gly Ser  
 805 810 815  
 Ala Thr Pro Arg Tyr Leu Tyr Leu Ala Ser Asn His Ser Asn Lys Trp  
 820 825 830  
 Gly His Pro Arg Gly Tyr Arg Ile Gln Met Leu Ser Phe Ala Gly Glu  
 835 840 845  
 Pro Leu Pro Gln Asn Ser Ser Met Ala Arg Gly Phe Ser Trp Glu Arg  
 850 855 860  
 Tyr Gln Leu Ala Val Thr Gln Arg Lys Glu Glu Glu Pro Ser Ser Ser

[illegible]